

AIを用いた橋梁床版の 補修要否・補修規模予測モデルの試作

(国研) 土木研究所 寒地土木研究所 寒地構造チーム ○角間 恒
佐々木 優太
西村 敦史

本研究では、寒冷地の鋼橋コンクリート床版を対象に、橋梁諸元や供用環境、点検結果等を統合したデータセットを用い、機械学習手法による床版上面損傷の補修要否予測モデルを構築した。その上で、複数のデータ分割およびハイパーパラメータ設定を考慮してモデル性能の安定性や未知データに対する予測特性を評価するとともに、SHAP値による損傷要因分析を行った。さらに、補修面積率を指標とした補修規模予測についても試行的に検討した。

キーワード：橋梁床版、補修要否、補修規模、AI予測

1. はじめに

橋梁の老朽化が進む我が国においては、人材や予算に制約がある中で、より効率的な維持管理の実施が求められている。コンクリート床版（以下、床版）では、床版上面に発生する土砂化が耐久性・供用性に影響を及ぼすものの、舗装裏で進行するため点検や調査による把握が容易ではなく、損傷把握の効率化が課題となっている。

近年、インフラ維持管理分野では、ビッグデータやAIを活用した予測・評価技術の適用が進み¹⁾、点検・調査や補修判断への活用が期待されている。

著者らは既報²⁾において、寒冷地の床版を対象に、橋梁諸元、供用環境、補修履歴、定期点検結果等を統合したデータセットを作成し、機械学習手法の一つであるLightGBMを用いて、床版上面コンクリートの損傷予測モデルを構築するとともに、その基本的な予測性能を確認した。

本研究では、これらの検討を継続・発展させるものとして、LightGBMによる補修要否予測に対し、複数のデータ分割およびハイパーパラメータ設定を考慮したモデル性能の安定性について評価を行う。また、未知データに対する予測精度の検証や、SHAP値を用いた床版上面損傷要因の分析を実施する。さらに、回帰モデルおよび分位点回帰モデルを用いた補修規模（補修面積率）予測についても試行的に検討し、その予測結果の傾向を整理する。

2. 学習用データセット

本章では、補修要否予測モデル（第3章）および補修規模予測モデル（第4章）の構築に用いた学習用データ

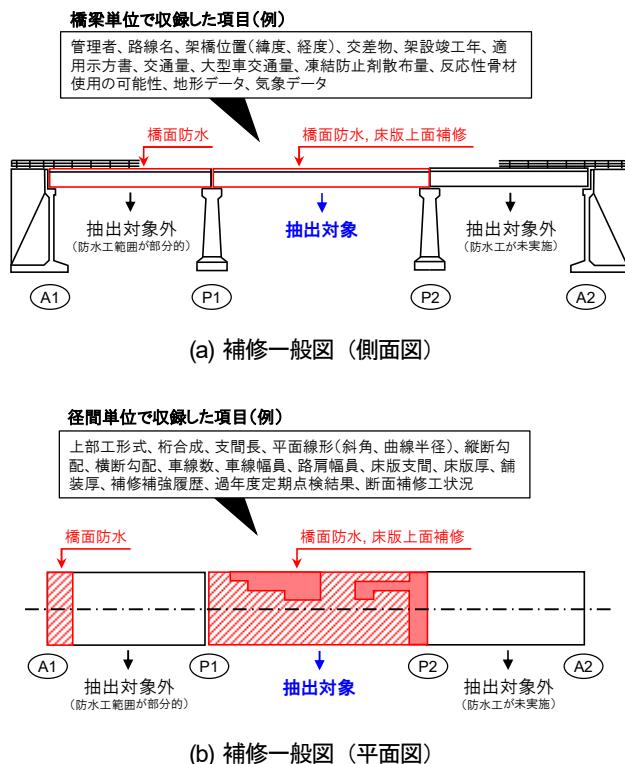


図-1 データセットの概要

セットの概要を説明する。データセットの基本的な構成および作成方法は既報²⁾で示したものと同一であり、本研究では既報のデータセットに新たなサンプルを追加している。データセットの詳細については、同文献を参照されたい。

(1) 概要

データセットは、国土交通省北海道開発局が管理する

表-1 データセットに収録した変数の一覧

変数区分	項目名	変数種別	変数名 ^{※1,※2}
説明変数	橋梁諸元	数値	経過年数*、支間長*、曲率* (=1/曲線半径)、斜角*、縦断勾配*、横断勾配、車線数、車線幅員合計*、路肩幅員合計*、床版支間*、床版厚*、舗装厚*
		カテゴリ	分割区分*(上下一体、上下分離)、交差物*、適用示方書、上部工形式、桁合成*、横断勾配種類*
	供用環境	数値	交通量、大型車交通量*、凍結防止剤散布量*、平均標高、最大傾斜角度*、合計降水量(年間*、春、夏、秋、冬)、日最高気温(年平均、春、夏*、秋、冬)、日最低気温(年平均、春、夏、秋、冬*)、日平均気温(年平均、春、夏、秋、冬)、合計日照時間(年間、春、夏、秋、冬)、全天日射量(年間*、春、夏、秋、冬)
		カテゴリ	反応性骨材使用の可能性*
	補修・補強履歴	カテゴリ	橋面防水工履歴*、断面修復工履歴*、床版補強履歴*
	定期点検結果	カテゴリ	剥離・鉄筋露出の対策区分判定*、床版ひびわれの対策区分判定、漏水・遊離石灰の対策区分判定*
目的変数	断面修復工	数値	断面修復工の面積、断面修復工の体積
		カテゴリ	断面修復工の有無*

※1:「*」印を付した変数は、予測モデルの構築(第3章、第4章)に使用した変数を示す。

※2:気象データにおいて、3~5月を春、6~8月を夏、9~11月を秋、12~2月を冬に位置付けた。

道路橋のうち、コンクリート床版を有する鋼鉄桁橋または鋼H桁橋を対象とし、1径間を1サンプルとして構成している(図-1)。対象とした径間は、平成26年度から令和6年度にかけて、橋梁補修工事内で、既設舗装(表基層)の撤去を伴う橋面防水工が実施されたものである。ただし、以下のいずれかに該当する径間はデータセットの対象外とした。

- 橋面防水工の施工延長が1径間未満または施工幅が1車線未満の径間
- 橋面防水工とともに床版の全層打換えが実施された径間
- 過去の補修が多岐にわたるなどにより補修履歴の追跡が困難な径間

これらの条件を満たした径間を抽出した結果、令和7年12月末時点でのサンプルサイズは433径間(190橋)となった。

(2) データセット項目

データセットでは、全サンプルに対して、橋梁諸元、供用環境、補修・補強履歴、定期点検結果の情報を説明変数として整理した。また、目的変数として、橋面防水工とともに床版上面コンクリートの断面修復工が実施されたか否か、およびその実施数量を整理した。各情報は、橋梁管理カルテ、定期点検調書、補修補強工事調書、および補修工事図面の記載内容に基づいている。

表-1にデータセットに含まれる全変数の一覧を示す。このうち、第3章および第4章で使用した変数には、表中で「*」印を付した。

3. 補修要否予測モデル

本章では、床版上面の補修要否を予測する機械学習モデルについて、予測性能の評価および損傷要因の分析を行う。本稿では、「断面修復工の有無」を目的変数とした二値分類モデルを用いる。

なお、基本的なモデル構築手法は既報²⁾と同様であるが、既報では单一条件下での性能評価を行っていたのに対し、本研究では複数ラン・複数トライアルによる学習と検証を実施し、再現性を考慮した汎化性能評価を行った。

(1) 機械学習モデルの構築手法

a) 機械学習手法および実装環境

補修要否予測モデルは、機械学習手法の一つである勾配ブースティング決定木(Gradient Boosting Decision Tree、以下、GBDT)³⁾により構築した。GBDTは、複数の決定木を逐次的に作成し、誤差を段階的に修正することで予測精度を高めるアンサンブル学習手法である。本検討では、GBDTのフレームワークの中でも計算効率や予測精度に優れるLightGBM⁴⁾を用いた。モデルはPython 3.12を用いて実装し、主なライブラリとしてpandas(ver. 2.2.2)、numpy(ver. 1.26.4)、scikit-learn(ver. 1.6.1)、lightgbm(ver. 4.5.0)、optuna integration(ver. 4.3.0)、imbalanced-learn(ver. 0.12.3)、shap(ver. 0.47.2)、japanize-matplotlib(ver. 1.1.3)を使用した。

b) 説明変数および前処理

表-1中に、モデル構築に使用した変数を「*」印で示す。説明変数については、相関の高い変数同士がモ

ルの解釈性や安定性に悪影響を与える可能性があるため、相關クラスタリングを実施したうえで、平均相関値および既往の床版損傷に関する技術的知見を考慮して選択した。また、カテゴリ変数にはダミーエンコーディングを適用した。

目的変数のクラス分布に偏りがある場合、学習データに対して SMOTE⁹を用いたオーバーサンプリングを実施し、少數クラスのサンプルを合成することで分布の不均衡を緩和した。

c) モデル構築

モデルの評価には交差検定を用い、分割数は 10 とした。本研究で用いたデータセットは全 433 サンプルから構成されており、このうち、平成 26 年度から令和 5 年度までの橋梁補修工事に基づく 409 サンプルを学習および交差検定に使用し、令和 6 年度の橋梁補修工事に基づく 24 サンプルを未知データとしてモデルの汎化性能評価に用いた。データ分割にあたっては、層化グループ抽出を用いることで、学習データと検証データにおける目的変数のクラス分布が偏らないようにするとともに、同一橋梁の径間が両データに分割されないよう配慮し、データリーアクを防止した。

さらに、予測精度の向上を目的として、LightGBM の主要ハイパーパラメータについて、自動探索ツールである Optuna¹⁰を用い、学習データに対する交差検定結果に基づく F1 値が最大となるように最適化を行った。

なお、既報²では、単一のデータ分割かつ交差検定内でハイパーパラメータを最適化していたため、予測精度が過大に評価されていた可能性がある。そこで、本研究では、データ分割およびハイパーパラメータ設定の違いによるモデル性能のばらつきを評価するため、データ分割を変更した条件（以下、ラン）と、各ラン内で異なるハイパーパラメータ設定を与えた条件（以下、トライアル）を設定し、複数ラン・複数トライアルに基づくモデル構築および性能評価を行った。具体的には、5 ラン × 10 トライアルの条件で検討を行っている。

また、Optuna によるハイパーパラメータ探索は、最終的なモデルの性能評価とは独立に実施し、単一のデータ分割に対して複数のトライアルおよび交差検定を行うことで最適解を決定した。併せて、LightGBM の出力する予測確率を二値の判定結果に変換するための判定閾値（threshold）についても最適化を行った。

d) 性能評価

モデルの性能評価には、適合率、再現率、および F1 値を指標として用いた。最終的なモデル性能は、複数ランおよび複数トライアルにおける評価結果を総合的に整理することで、特定のデータ分割やハイパーパラメータ設定に依存しない、汎化性能を考慮した評価とした。

また、各説明変数が予測結果に与える影響を把握するため、SHAP 値¹¹（Shapley Additive exPlanations）を算出し、予測への寄与度を可視化した。

表-2 予測結果の混同行列（最良トライアル）

		正解	
		補修なし	補修あり
予測	補修なし	137 (=TN)	8 (=FN)
	補修あり	101 (=FP)	163 (=TP)

表-3 予測モデルの性能評価指標（平均±標準偏差）

モデル	適合率	再現率	F1 値
最良ラン	0.592 ±0.021	0.923 ±0.019	0.707 ±0.016
全体モデル	0.601 ±0.021	0.867 ±0.056	0.691 ±0.024

適合率 = $TP/(TP+FP)$ ：「補修あり」と予測したもののうち、実際に「補修あり」であった割合（=誤検出の少なさ）

再現率 = $TP/(TP+FN)$ ：実際に「補修あり」であったもののうち、正しく「補修あり」と予測できた割合（=見逃しの少なさ）

F1 値 = $2TP/(2TP+FP+FN)$ ：適合率と再現率の調和平均（=総合的なモデル精度）

（2）判定閾値（threshold）の最適化

1 ラン × 10 トライアルにおいて得られた最適な判定閾値は、概ね 0.3～0.4 の範囲に分布し、平均値は 0.356 であった。この結果を踏まえ、本研究では、以降のモデル構築および性能評価において、判定閾値として threshold=0.35 を採用した。

（3）交差検定による性能評価

表-2 に、予測モデルの性能評価結果の一例として、最良ラン（10 トライアルの F1 値平均が最大となったラン）に含まれるトライアルのうち、F1 値が最大となったトライアル（以下、最良トライアル）について、交差検定結果を統合した混同行列を示す。当該トライアルにおいて、「補修あり」 171 件のうち 163 件が正しく予測され、8 件が誤分類された。また、「補修なし」 238 件のうち 137 件が正しく予測され、101 件が誤分類された。

表-3 に、予測モデルの性能評価結果を示す。表中の「最良モデル」は前述の最良ランに基づく評価結果を、「全体モデル」は計 50 トライアル（5 ラン × 10 トライアル）の評価結果を統合した平均性能を示す。表に示すとおり、最良モデルおよび全体モデルのいずれにおいても再現率が高く、補修が必要な床版を見逃しにくい予測性能が得られた。特に最良モデルでは再現率が 0.923 と高く、補修対象を適切に抽出できていることが確認できる。一方で適合率はやや低い値を示しているが、これは補修不要な床版を一部安全側に判定していることを意味しており、見逃しを抑制する傾向を有する特性であるといえる。

また、各指標の標準偏差はいずれも小さく、データ分

表4 未知データに対する予測結果の混同行列

		正解	
		補修なし	補修あり
予測	補修なし	5(=TN)	0(=FN)
	補修あり	10(=FP)	9(=TP)

表5 未知データに対する予測モデルの性能評価指標

モデル	適合率	再現率	F1 値
最良トライアル	0.474	1.000	0.643

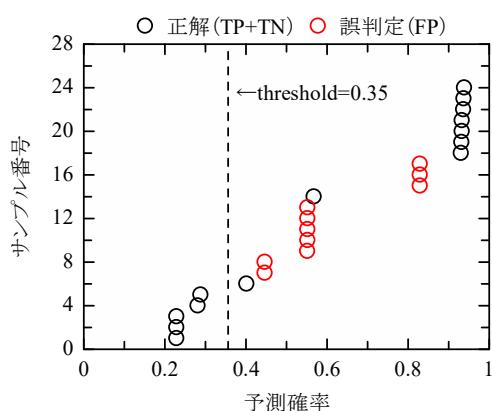


図2 予測確率の分布

割やハイパーパラメータ設定の違いによる性能のばらつきが小さい、安定したモデルであることが確認できた。

なお、既存モデル²⁾における適合率、再現率、F1 値はそれぞれ 0.747、0.813、0.779 であったのに対し、本研究における全体モデルでは、再現率を維持しつつ、より安定した性能評価が得られた。一方で、適合率および F1 値は既存モデルに比べて低下したように見えるが、これは単一のデータ分割および交差検定内でハイパーパラメータを最適化していた従来手法に対し、本研究では複数のデータ分割およびハイパーパラメータ設定に基づいて性能評価を行った結果、特定条件に適合した性能が平均化されたためと考えられる。すなわち、これらの結果はモデル性能の低下を示すものではなく、汎化性能をより適切に反映した評価である。

(4) 未知データによる検証

次に、モデル構築に使用していない未知データ（24 サンプル）を用いて、予測モデルの汎化性能を検証した。なお、本検証には、前述の最良ランに含まれる最良トライアルモデルを使用した。

表4 に未知データに対する混同行列を、表5 に性能評価指標を示す。その結果、補修が必要な 9 件は全て正しく予測され、見逃し（FN）は発生しなかった。一方で、補修不要な 15 件のうち 10 件を補修ありと判定しており、適合率は 0.474 となった。以上より、本モデルは

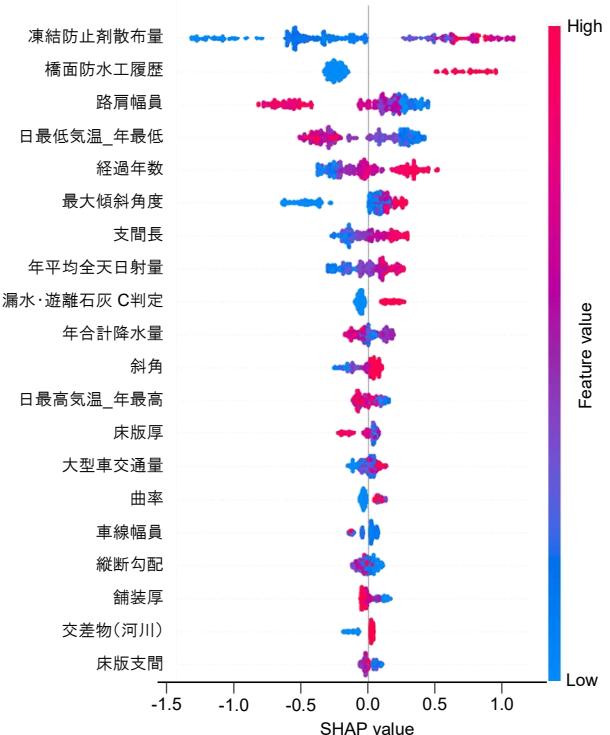


図3 各説明変数の SHAP 値（上位 20 変数）

補修が必要な床版を確実に抽出できる一方で、判定は安全側となる傾向があることが確認できた。維持管理の観点からは、補修が必要な床版の見逃しを防ぐことが重要であり、本モデルの特性は、補修候補の一次抽出等の場面において有用となる可能性がある。

図2 に未知データに対する予測確率の分布を示す。誤判定となったサンプル（赤色の点）は、主に予測確率 0.5 附近で見られるが、予測確率が高いにも関わらず誤判定となったサンプルも一部存在する。このことから、判定閾値による二値分類に加えて、予測確率に基づいて補修リスクを段階的に評価する方法（例：高・中・低リスク）を併用することで、実務において本モデルのより柔軟な活用が可能になると考えられる。

(5) SHAP 値による損傷要因の分析

本項では、SHAP 値を用いて各変数が予測結果に与える影響を可視化し、床版上面の損傷発生要因について分析する。

図3 に、SHAP 値による変数の影響度を示す。本図では、最良ランにおける計 10 トライアルの予測結果から得られた SHAP 値の平均を用いている。横軸は SHAP 値の大きさ、縦軸は影響度の大きい上位 20 変数を示しており、色は各変数値の大小を表す。また、SHAP 値の符号は、正が損傷リスクを高める方向、負が低減する方向への寄与を示す。

予測モデルに用いた説明変数のうち、最も影響が大きいのは「凍結防止剤散布量」であり、次いで「橋面防水工履歴」「路肩幅員」「日最低気温の年最低値」「経過

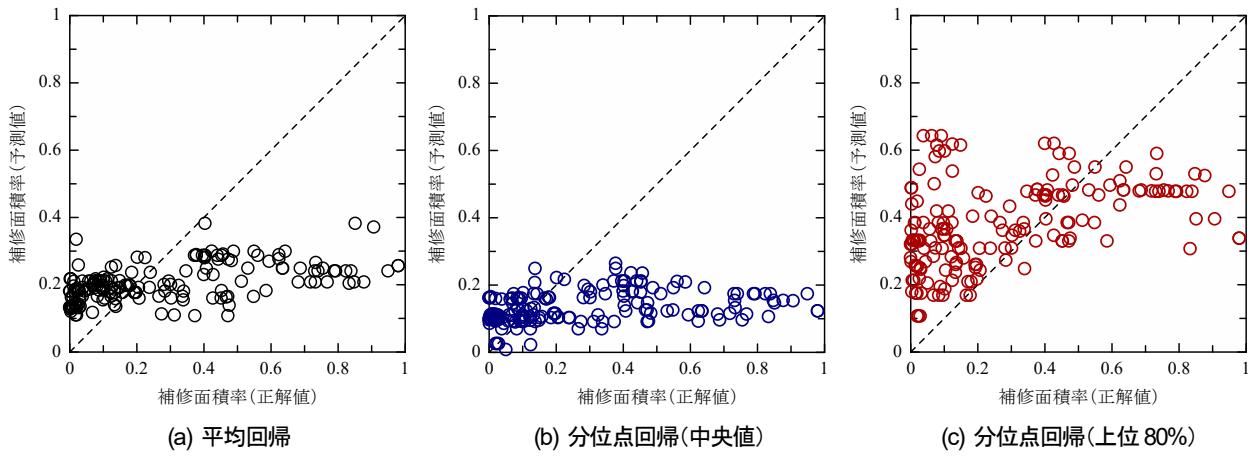


図4 回帰モデルによる予測結果

年数」と続いた。凍結防止剤散布量、気温条件、経過年数といった上位変数は、床版上面の損傷発生と関係が深い要因として、維持管理の現場においても広く認識されているものである。

一方、橋面防水工履歴については、実務上の一般的な認識とは反対に、履歴ありの場合に損傷リスクが高まる方向の影響が確認された。この傾向は、既報²⁾のモデルにおいても同様に示されており、本研究においてもその再現性が確認された。橋面防水工履歴および路肩幅員と床版上面損傷との関係についての解釈は既報で考察しており、本稿では詳述しない。

さらに、最大傾斜角度、支間長（径間長）、床版厚、床版支間といった構造条件、日射量、降水量、気温条件などの環境条件、ならびに大型車交通量、曲線半径、車線幅員、縦断勾配、舗装厚、交差物件（河川交差）といった交通・路線条件に関わる変数が、影響度の高い要因として抽出された。これらは、床版上面の損傷発生が単一の要因ではなく、複数の条件が組み合わさって生じていることを示している。

以上のように、SHAP 値により抽出された影響度の高い変数には、経験的に重要と認識してきた要因に加え、点検や調査の中では必ずしも明示的に整理されてこなかった要因も含まれている。一方で、一般に床版損傷との関係が強いと想定してきた要因（例えば大型車交通量）であっても、本モデルでは相対的に影響度が低い結果となった変数も確認された。

これらの結果から、本研究で構築した AI モデルは、補修要否を判定するための手法にとどまらず、SHAP 値を通じて、点検・調査時の着目点を整理するためのツールとしても有用になる可能性がある。

4. 補修規模予測モデル

本章では、補修が必要と判定された床版を対象に、補修規模の予測を試みる。補修規模を推定する回帰モデル

表6 平均回帰モデルの性能評価指標（平均土標準偏差）

モデル	RMSE	MAE
最良モデル	0.238±0.003	0.197±0.003
全体モデル	0.243±0.006	0.198±0.005

および分位点回帰モデルを用い、予測結果の傾向を簡単に整理する。なお、本検討における補修規模予測モデルは、現時点では試行的な検討段階にあり、予測結果の傾向把握を主眼としている。

(1) モデル概要

補修規模予測モデルは、補修要否予測モデルと同一の説明変数を用い、補修規模を連続値として予測する回帰モデルとして構築した。学習および検証には、データセットのうち平成 26 年度から令和 5 年度までの橋梁補修工事に基づき、「補修あり」と整理された 171 サンプルのみを使用した。目的変数としては、橋面（径間）全体面積に対する補修面積の割合（以下、補修面積率）を設定した。

モデルには LightGBM を用い、補修面積率の平均値を予測する通常の回帰モデル（以下、平均回帰モデル）に加え、予測値のばらつきを考慮するため、分位点回帰モデル（50%分位、上位 80%分位）を併せて検討した。分位点回帰モデルは、目的変数の条件付き分布における特定の分位点（例えば中央値や上位分位）を直接推定する手法であり、平均値予測では捉えにくい分布の偏りや不確実性を考慮できる点に特徴がある。

ハイパーパラメータの最適化には Optuna を用い、平均回帰モデルでは RMSE が最小となるように、分位点回帰モデルでは損失関数（quantile loss）が最小となるように探索を行った。

(2) 予測結果

図4 に、平均回帰モデルおよび分位点回帰モデルに

による補修面積率の予測結果を示す。

a) 平均回帰モデル

平均回帰モデルでは、個々のサンプルに対する正解値と予測値のばらつきが大きく、本データ条件下において補修規模を精度良く推定するには至っていない。予測値は、学習データ全体での補修面積率の平均値（0.28）および中央値（0.18）付近に集中する傾向があり、補修面積率分布のばらつきや極端値を十分に捉えにくいことが確認された。なお、RMSE および MAE による定量評価結果（表-6）からも、最良モデルと全体モデルの性能差は小さく、回帰モデルとしての予測精度には一定の限界があることがわかる。

b) 分位点回帰モデル

次に、分位点回帰モデル（中央値、50%分位）では、補修面積率が小さいサンプルに対して、平均回帰モデルと比較して正解値に近い予測が得られる傾向が確認された。一方で、補修面積率が大きいサンプルに対しては過小に予測される傾向があり、大規模補修の定量的な推定には課題が残る。

これに対し、分位点回帰モデル（上位 80%分位）では、補修面積率が大きいサンプルに対して高めの予測値を出力する傾向が見られ、平均回帰モデルや中央値モデルと比べて、補修面積率が大きくなる可能性を考慮した予測が可能となった。特に、サンプル数の少ない大規模補修の領域において、過小評価を抑制しようとする特性が確認された。

以上より、平均回帰モデルは補修規模の平均的・中央的な傾向を、中央値モデルは比較的小規模な補修の傾向を、上位分位モデルは補修規模が大きくなる可能性を示す指標として、それぞれ異なる役割を持つことがわかった。一方で、本検討における補修規模予測は、現時点では補修規模を定量的に予測する段階には至っていない。これは、大規模補修サンプルの不足や説明変数の制約に起因する予測の偏りの影響が大きいと考えられ、今後は学習データの拡充や新たな説明変数の追加等を通じて、補修規模推定精度の向上を図る必要がある。

5. おわりに

本研究では、寒冷地の鋼橋コンクリート床版を対象として、橋梁諸元、供用環境、点検結果等を統合したデータセットを用い、機械学習手法による補修要否および補

修規模の予測について検討を行った。

その結果、補修要否予測では、学習条件によらず一定のレベルの予測性能が安定して得られることが確認された。特に再現率が高く、補修が必要な床版を見逃しにくい特性を有していることから、補修要否検討の補助ツールとして、現時点でも実務への適用が可能な水準にあると考えられる。また、SHAP 値を用いた説明変数の影響度の可視化により、点検・調査時に着目すべき項目の整理にも活用できることが示された。

一方、補修規模予測（補修面積率の推定）については、補修面積率のばらつきや大規模補修サンプルの少なさ等の影響を受け、現時点では精度の高い予測は困難であることが明らかとなった。

補修要否予測および補修規模予測に共通する今後の課題として、学習データの拡充や説明変数の充実、モデルの改良等を通じた、予測精度および実務への適用性の向上が挙げられる。

謝辞：本検討の実施に当たり、国土交通省北海道開発局から橋梁の維持管理に関する各種資料をご提供いただきました。ここに付記し、関係各位に感謝の意を表します。

参考文献

- 1) 土木学会 SIP インフラ連携委員会：インフラ維持管理への AI 技術適用のための調査研究報告書、2019.
- 2) 角間恒、西村敦史：寒冷地の橋梁床版を対象とした損傷予測 AI の構築と基本性能評価、寒地土木研究所月報、No.876, pp.39-43、2025.
- 3) Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, Annals of Statistics, Vol.29, No.5, pp.1189-1232, 2001.
- 4) Ke, G. et al.: A Highly Efficient Gradient Boosting Decision Tree, Advances in Neural Information Processing Systems 30, pp.3146-3154, 2017.
- 5) Chawla, N.V. et al.: Synthetic Minority Over-sampling Technique, Journal of Artificial Intelligence Research, Vol.16, pp.321-357, 2002.
- 6) Akiba, T. et al.: A Next-generation Hyperparameter Optimization Framework, Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2019), pp.2623-2631, 2019.
- 7) Lundberg, S.M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, Advances in Neural Information Processing Systems 30, pp.4765-4774, 2017.